DOCUMENT RESUME

ED 137 296 SP 010 947

AUTHOR Nicolich, Mark J.; Weinstein, Carol S.
TITLE Time Series Analysis of Behavioral Changes in an Open Classroom.
PUB DATE 77
NOTE 14p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Analysis of Variance; *Behavioral Sciences; *Classroom Observation Techniques; Educational Research; Factor Analysis; Research Criteria; Research Design; *Research Methodology; *Research Problems; *Statistical Analysis; Statistical Data
IDENTIFIERS *Time Series Analysis

ABSTRACT

This paper discusses the problems of gathering and analyzing data on classroom behaviors. The difficulties involved in using the common pretest-posttest experimental design are described. To grapple with these problems, a study was conducted on the effect of environmental changes on student behavior, employing a time-series design and time-series analysis. The hypothesis tested was that minor changes in the physical setting of the classroom could produce predictable, desirable changes in students' behavior. Formal observations of students' locations and activities were conducted six times a day for two weeks, using a time-sampling-by-child instrument. The data from the six daily observations of all children were reduced to daily summary statistics indicating the percentage of time children were observed in a particular area or engaged in a specific activity. The data were then analyzed using time-series techniques. Results indicated that most of the desired behavior changes were produced, thus supporting the hypothesis. The same experiment was conducted using the traditional statistical methods. The conclusion was reached that time series techniques appear to have great potential for educational research. (JD)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED137296 *

Time Series Analysis of Behavioral Changes in an Open Classroom

Mark J. Nicolich and Carol S. Weinstein

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

One of the most commonly used research paradigms in behavioral science is the pre-test--post-test experimental design involving independent subjects divided into experimental and control groups. The data collected in a study like this are typically analyzed with analysis of variance, t-test or chi square techniques. Such analyses are straightforward and valid if all the assumptions of the chosen technique have been met. Of particular importance is the requirement that the observations be conducted on independent subjects.

This requirement poses a knotty problem for the educational researcher who is observing the behavior patterns of interacting students. In this case an obvious dependency exists among subjects. In addition, the difficulty of perfectly matching classrooms makes the selection of a control group particularly troublesome. Since a pre-test--post-test design <u>without</u> a control group is hardly adequate, observations of the students can be conducted at <u>several</u> time periods before and after treatment. This would transform the pre-test--post-test situation into a quasi-experimental time-series design, which would be better able to distinguish changes due to the treatment from those caused by a steady trend (brought about by history or maturation, for example) or by experimentally unplanned classroom events.

Successive observations of the same subjects are clearly not independent, however, and do not fall into any class of ANOVA techniques which could normally be used to assess a treatment effect. Indeed, Scheffé (1959),

SP010 949

discussing the effects of departures from the underlying assumptions of ANOVA on error rates and estimation has stated: "In general, of the three kinds of possible departures from assumptions we have considered [normality of errors and effects, equality of variance of the errors, and statistical independence of the errors], those caused by lack of independence are the most formidable to cope with" (p. 364).

Research on the relationship between environment and behavior in general--and classroom settings and student behavior in particular--is often plagued by problems of dependency among subjects and over time. Perhaps for this reason, several studies of man-environment relations have been purely descriptive, with no statistical analysis: Brunetti (1972); Hahn (1973); Rivlin and Wolfe (1972); and Zifferblatt (1972). There have also been a number of studies which ignore the independence requirement and use inappropriate or questionable statistical methods, such as chi square, t-tests, or analysis of variance: Brookes and Kaplan (1972); Coates and Sanoff (1972); McGrew (1970); and Winett, Battersby, and Edwards (1975).

In an attempt to grapple with these problems, a study by Weinstein (1975), on the effect of environmental changes on student behavior, employed a time-series design and time-series analysis. The experimenter observed the spatial distribution of activity in a second-third grade informal classroom divided into various interest areas, before and after a change in room arrangement. She tested the general hypothesis that minor changes in the physical setting could produce predictable, desirable changes in students' behavior. Formal observations of students' locations and activities were conducted six times a day for two weeks, using a

time-sampling-by-child instrument. Data were recorded on a floor plan of the room. A descriptive analysis of this pre-change data revealed several problems with the way the room and the materials in it were being used (e.g. certain areas were crowded, while others were under-used; students were not involved with the manipulative materials to the extent the teacher desired). Design changes were then made with specific behavioral goals in mind, and the two-week post-change observation period was begun.

The widely used statistical techniques--ANOVA, t-test and χ^2 were found to be inadequate to assess the effect of the design changes. Serial dependency among observations for one subject negated the use of the t-test, and the variation in correlation among the series of observations negated use of repeated measure ANOVA. The chi square test could be used if it was assumed that entries in an area population-by-pre-post contingency table were a sample of time periods (thus achieving independence of the observations, which could not be assumed if the observations were a sample of subjects). Still, the chi square test could not distinguish between a change due to the intervention and a change due to a steady trend.

With these difficulties in mind, time-series analysis (Box and Jenkins, 1970) appeared to be the most satisfactory solution. The data from the six daily observations of all children were reduced to daily summary statistics indicating the percentage of time children were observed in a particular area or engaged in a specific activity. The resultant series of pre- and post-change data points, one for each day, were then analyzed using time-series techniques. Results indicated that most of the desired behavior changes were produced, definitely supporting the general hypothesis that relatively simple changes in room arrangement can produce

specific, desired changes in behavior.

Since other investigators dealing with similar questions had used statistical techniques rejected by the experimenter, the present authors were interested in seeing what kinds of results would have been obtained had standard t-tests been used--in particular, the frequency of type I errors and estimates of the magnitude of the intervention effect. Three specific analytical techniques were considered. The first was the standard, independent, two-sample t-test assuming equal population variances, and estimating a shift by the difference of post-test and pre-test series means. The second method was the basic time series analysis technique promulgated by Box and Jenkins (1970), based on maximum likelihood estimation. The third method was a Bayesian time series analysis put forth by Box and Tiao (1965). To assess the difference among these analytical techniques, a series of data sets, based on the Weinstein situation, were generated. Each set consisted of 40 pre-test and 40 post-test observations. The basic model used in the Weinstein study and for the generation of these data sets was the integrated first order moving average model, which Box and Tiao (1965) recommend as a frequently encountered form of time-series models:

$$
\begin{aligned}
z_1 &= L+a_1 \\
z_t &= L+\gamma \sum_{j=1}^{t-1} a_j+a_t \qquad t=2,3,\ldots,m \\
z_t &= \delta+L+\gamma \sum_{j=1}^{t-1} a_j+a_t \qquad t=m+1,m+2,\ldots,n
\end{aligned}
\tag{1}
$$

$$0< \gamma <2$$

where $\{a_t\}$ is a series of independent random normal deviates with mean 0 and variance σ^2, L is the base level, δ is the change in level after time m--caused by the intervention--and γ is the proportion of past deviations incorporated into the present observation, also called the moving average parameter.

The model can be written in two other equivalent forms. Equation 2 indicates that the present observation is based on the previous observation, a portion of the previous deviation and a present deviation. Equation 3 indicates that the first differences form a series based on a present deviation plus a portion of the previous deviation.

$$
\begin{aligned}
z_1 &= L+a_1 \\
z_t &= z_{t-1} + (\gamma-1)a_{t-1}+a_t \qquad t=2,3,\ldots,m,m+2,\ldots,n \\
z_{m+1} &= \delta + z_m+(\gamma-1)a_m+a_{m+1} \\
&\qquad 0< \gamma <2
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
z_1 &= L+a_1 \\
\text{and} \\
(z_t-z_{t-1}) &= \gamma\, a_{t-1}+(a_t-a_{t-1}) \qquad t=2,3,\ldots,m,m+2,\ldots,n \\
(z_{m+1}-z_m) &= \delta + \gamma\, a_m+(a_{m+1}-a_m) \\
&\qquad 0< \gamma <2
\end{aligned}
\tag{3}
$$

In all, 90 sets of data were generated, and each set was analyzed by each of the three techniques. The standard deviation of the a_t took on values of .1, .5 and .7; δ took values of 0, .5, 1.0, 1.5 and 2.0; and L was 10 for all cases. The values of γ determine the dependency among the terms in the series. When γ is 1 there is no serial correlation among the first differences of the observations; that is, the present observation is the previous observation plus the random component a_t (see equation (2)). When γ is zero the present observation is a uniformly weighted sum of previous observations which is an extreme correlational situation. In the simulations the values of γ were 0, .25, .50, .75 and 1.0. Values in the range 1.0 to 2.0 were not considered as they would represent a negative weighting of past observations which is not consistent with many behavioral models. One set of data was generated for each combination of σ, δ and γ yielding 75 data sets, and a second replication was done with $\gamma = 1$ for all combinations of σ and δ yielding an additional 15 data sets.

What happens is a brief outline of the time series techniques, beginning with the method of Box and Tiao (1965). The technique is to calculate the lagged correlations for the pre-test and post-test data. On the basis of these correlations a suitable model is chosen (if differencing of the observations is indicated, in order to remove nonstationarity, the lagged correlations are calculated from the differenced data). The model is composed of some combination of autoregressive, moving average and differencing terms. To analyze the simulated data, the form of the model was assumed to be known as a first order moving average on the first differences of the observations. If the parameter γ is unknown, it is not possible to directly estimate the value of δ. The approach is to employ a Bayesian

analysis using sample information about γ to make estimates of δ.

In the model (Eq. 1), there are four unknown parameters-- δ, L, γ, and σ. The distribution of δ conditional on γ and the observations can be calculated and plotted as a function of γ as γ ranges from 0 to 2. A similar conditional distribution can be calculated for the parameter L. Of major importance is the posterior distribution of γ assuming no prior information (a uniform prior distribution for γ) and the value of the t statistic to test the hypothesis δ=0, as γ varies from 0 to 2. For illustrative purposes, a plot of these functions is shown in Figure 1 for the data generated when δ = 0.5, σ = .5, and γ = .25. The Bayesian analysis based on the functions in Figure 1 indicates that, if there is no prior knowledge of γ, the posterior distribution, $h(\gamma/Z)$, has its support over the range of δ from .17 to .75. Over this entire range the t value is above the 1% critical value of 2.33, indicating a δ significantly different from zero. The estimated value for δ would be in the range of 1.62 to 2.01. The analysis clearly indicates a non zero value for δ and a γ in the range of .17 to .75.

Maximum likelihood analysis, on the other hand, would choose the value of γ which has the maximum probability of the posterior distribution of γ. In this case it is at $\hat{\gamma}$ = ~~.51~~ .39. At this point t = 3.76 and $\hat{\delta}$ = 1.88. From the other functions, not graphed, it is found that the maximum likelihood estimates of L and σ are 10.15 and .42 respectively. The estimation is accurate. The t-test estimated δ at .29 with a t value of 1.28, which lends to the false conclusion of no change in level.

There is little difference in the final result (relative to the estimation of δ with a uniform prior) between the two time series technqiues. In the remaining discussion, the acceptance or rejection of the hypothesis that $\delta = 0$ will apply to both the Bayesian and maximum likelihood analysis, and the estimate of δ and t value will be the maximum likelihood values.

The results of the analyses of the 90 simulations clearly demonstrated the advantages of using time series techniques for analyzing this type of data. In one of the six sets of simulated data with no serial correlation and $\delta = 0$, the t-test and time series tests both indicated a δ statistically significantly different from zero. The t value was 2.15 (1% level), but the estimate of δ was small ($\delta = -.06$). This can be attributed to an expected Type I error occurrence.

In the twelve simulations with $\delta = 0$ and serial correlation ($\gamma \neq 1$), the t-test analysis indicated a δ significantly different from zero in 11 cases; the mean of the absolute value of the t statistic was 5.44; the minimum value was 2.15. This corresponds to a Type I error rate of 92% when the testing was carried out at the nominal 5% level. The estimated magnitude of the shift was 1.39 with individual estimates ranging from 4.56 to -1.25. The time series techniques indicated two of the twelve simulations had statistically significant estimates of δ. When tests are carried out at the 5% level of significance, it would be expected that in 12% of the cases at least two significance values would be observed in 12 trials. Thus, two is not an extreme number of false rejections. It should be noted that both of these errors occurred at the maximum σ value and at γ values of 0 and .25. Since these represent the two most extreme trials, it appears that with highly variable or strongly correlated data the error rates may not be at their stated significance levels in this type of time series analysis.

For the 16 cases with no serial dependence and $\delta \neq 0$, the t-test and time series techniques yielded results similar to each other. Both indicated the presence of the shift and estimated its magnitude to a desirable degree of accuracy (the magnitude of the error increasing with σ). The remaining 48 cases of serial dependence and $\delta \neq 0$ indicated that the t-test was seriously in error. The problem was not so much in the Type II error rate, since there were only four non-rejections in 48 cases when the alternative hypothesis was two-sided, and 14 non-rejections for the appropriate one-sided hypothesis. However, the derived estimated value of δ is poor. For the 48 cases, the mean error of the estimate of δ was 21.8 standard deviation units, where the s.d. unit is $\left|\frac{\hat{\delta}-\delta}{\sigma/\sqrt{n}}\right|$. The corresponding error in δ for the maximum likelihood time series estimate is 6.9. The Bayesian time series analysis failed to reject the appropriate one-sided null hypothesis in only 9 of the 48 cases.

In summary, when serial correlation is present, both the t-test and time series techniques applied to this data have approximately the same Type II error rates (29% and 19%), but the estimate of the shift produced by the t-test is less accurate. The t-test has a Type I error rate near 100%, when tests are performed at the nominal 5% level, while the time-series technique has a Type I error rate that is not statistically significantly different from 5%. When no serial correlation is present all techniques perform equally.

It is important to point out that this analysis was carried out knowing the form of the model exactly, and only estimating the parameters. In actual situations, the form of the model must also be estimated from the

sample correlation coefficients. The additional estimations of a model could change the above stated results.

Time series techniques appear to have great potential for educational research. Experimenters attempting to fulfill the assumptions of the time series model will invariably produce more rigorously designed studies (e.g. by ensuring that they have a sufficient number of data points, that the intervals between observations are equal, etc.). In addition, awareness of these statistical procedures and their advantages will hopefully eliminate the use of inappropriate statistical methods and will stimulate research into areas which have previously been ignored due to the lack of appropriate statistical tools.

REFERENCES

Box, G.E.P., & Jenkins, G.M. Time series analysis, forecasting and control. San Francisco: Holden-Day, Inc., 1970.

Box, G.E.P., & Tiao, G.C. A change in level of a nonstationary time series. Biometrika, 1965, 52, 181-192.

Brookes, M.J., & Kaplan, A. The office environment: Space planning and affective behavior. Human Factors, 1972, 14(5), 373-391.

Brunetti, F.A. Noise, distraction and privacy in conventional and open school environments. Proceedings of the Environmental Design and Research Association Conference. Los Angeles: University of California, 1972.

Coates, G., & Sanoff, H. Behavioral mapping: The ecology of child behavior in a planned residential setting. Proceedings of the Environmental Design and Research Association Conference. Los Angeles: University of California, 1972.

Hahn, R.R. Behavioral evaluation of a juvenile treatment center: Case study of a planning methodology. In W.F.E. Preiser (Ed.), Environmental Design Research, (Vol. I). Stroudsburg, Pennsylvania: Dowden, Hutchinson, and Ross, Inc., 1973.

McGrew, P.L. Social and spatial density effects on spacing behavior in preschool children. Journal of Child Psychology and Psychiatry, 1970, 2, 197-205.

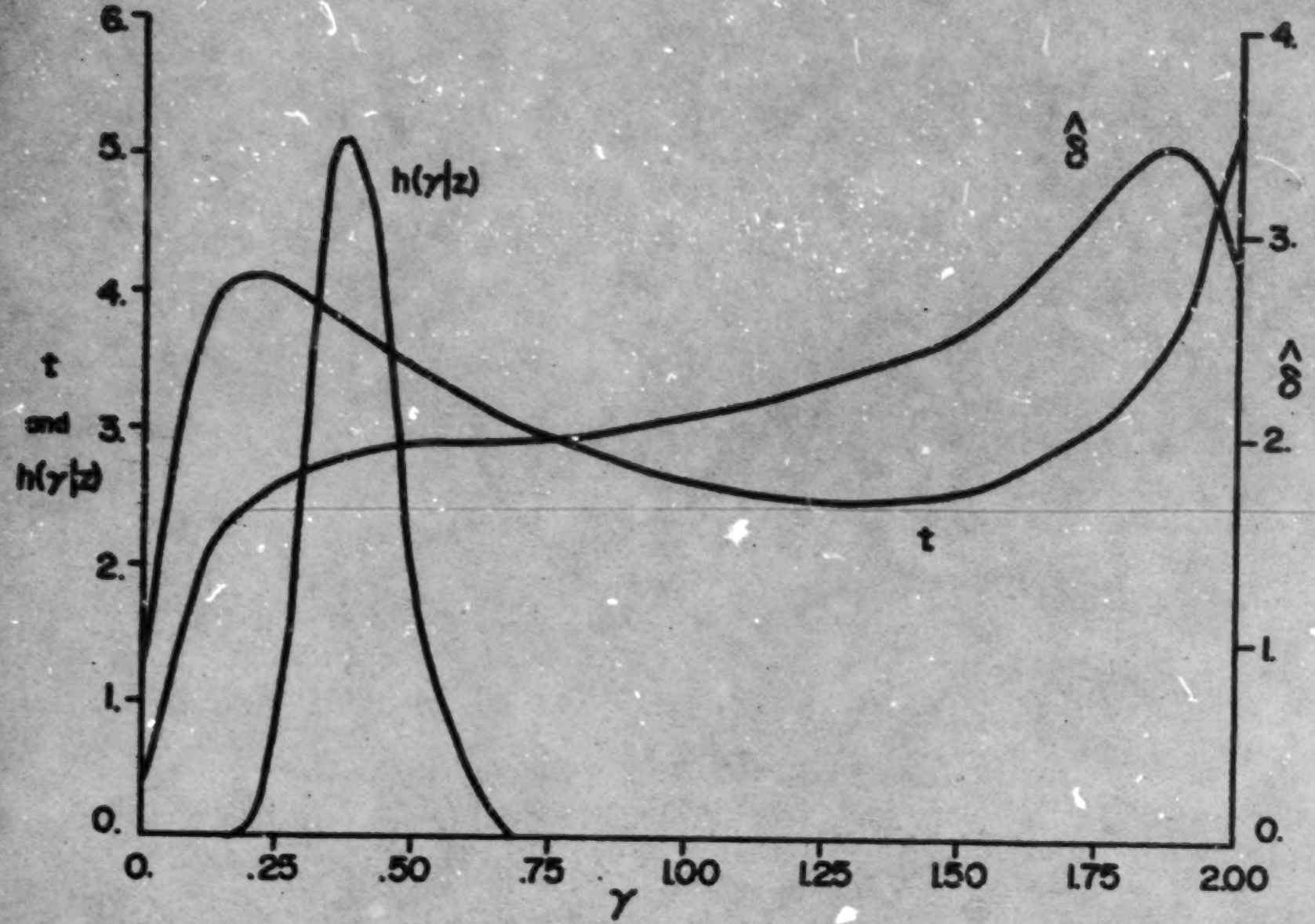
Rivlin, L., & Wolfe, M. The early history of a psychiatric hospital for children: Expectations and reality. Environment and Behavior, 1972, 4(1), 33-72.

Scheffé, H. The analysis of variance. New York: John Wiley and Sons, Inc., 1959.

Weinstein, C.S. The effect of a change in the physical design of an open classroom on student behavior. Unpublished thesis, Harvard University, 1975.

Winett, R. A., Battersby, C.D., & Edwards, S.M. The effects of architectural change, individualized instruction, and group contingencies on the academic performance and social behavior of sixth graders. Journal of School Psychology, 1975, 13(1), 28-40.

Zifferblatt, S.M. Architecture and human behavior: Toward increased understanding of a functional relationship. Educational Technology, 1972, 12(8), 54-57.


6.
5.
4.
3.
2.
1.
0.
t
and
h(y|z)
h(y|z)
$\hat{\delta}$
t
4.
3.
2.
1.
0.
$\hat{\delta}$
0.
.25
.50
.75
1.00
1.25
1.50
1.75
2.00
γ